> The prability associated with '*The cat died'* is the joint probability of '*The'* occurring as the first word in a sentence, of '*cat'* occurring given that the preceding word was '*The'*, and the probability of '*died'* occurring given that the preceding word was '*cat'*. The basic data for such a model is thus observations about the relative frequency of various pairs of words (bigrams), (ibid: 139).

From the above illustration, one can easily conclude that such models could be applied to certain languages which syntax and word order play an important role in its syntax while other languages such as Arabic are freer in word order and such models are doomed to failure.

In the translation model, a generalization might assume that the probability of ***a*** appearing as the translation of word ***e*** depends on the predecessors of ***a*** and ***e***. It is obvious however that there are shortcomings to the bigram model as the probability declines sharply in an example such as *The big fat black cat* compared to *the big cat* where *the* exerts the same kind of effect while the probability of the sequence *big fat and black* is less.

2.8. Alignment

In addition to drawing on the probability theory, statistical MT developers consider alignment as the backbone of their domain. Alignment can be divided according to the size of the unit of language. It starts with the text as the largest unit until it goes down to the word.

Although it is not part of the translation process *per se*, text alignment is used to create lexical resources such as bilingual dictionaries and parallel grammars, which then improve the quality of MT. It is evident that the availability of parallel texts or bitexts which include the same content in several languages is of paramount importance. Such kind of texts can be found in parliamentary proceedings and official documents in countries with multiple official languages. Another source could be found in magazines and newspapers published in more